

On the behavior of diffusion processes with traps

M. Freidlin*, L. Koralov†, A. Wentzell‡

Abstract

We consider processes that coincide with a given diffusion process outside a finite collection of domains. In each of the domains, there is, additionally, a large drift directed towards the interior of the domain. We describe the limiting behavior of the processes as the magnitude of the drift tends to infinity, and thus the domains become trapping, with the time to exit the domains being exponentially large. In particular, in exponential time scales, metastable distributions between the trapping regions are considered.

2010 Mathematics Subject Classification Numbers: 60F10, 35J25, 47D07, 60J60.

Keywords: Metastable Distributions, Large Deviations, Exit Problem, Non-standard Boundary Problem.

1 Introduction

Let v be an infinitely smooth vector field on the d -dimensional torus \mathbb{T}^d . (General manifolds, whether compact or not, can also be considered, but we'll stick with the torus for the sake of simplicity of notations later on.) Consider the process $X_t^{x,\varepsilon}$ defined via

$$dX_t^{x,\varepsilon} = \frac{1}{\varepsilon} v(X_t^{x,\varepsilon}) dt + dW_t, \quad X_0^{x,\varepsilon} = x, \quad (1)$$

where ε is a small parameter and W is a d -dimensional Brownian motion. This process is an order-one perturbation of a strong deterministic flow, or, equivalently, a time-changed small random perturbation of the deterministic flow $\dot{x}(t) = v(x(t))$. We'll be interested in the large-time behavior of the process when $\varepsilon \downarrow 0$.

More precisely, assume at first that the collection of asymptotically stable limit sets of the unperturbed flow consists of n equilibrium points O_1, \dots, O_n . Let D_1, \dots, D_n denote the sets that are attracted to O_1, \dots, O_n , respectively. Intuitively, when ε is small and t is large,

*Dept of Mathematics, University of Maryland, College Park, MD 20742, mif@math.umd.edu

†Dept of Mathematics, University of Maryland, College Park, MD 20742, koralov@math.umd.edu

‡Dept of Mathematics, Tulane University, New Orleans, LA 70118, wentzell@math.tulane.edu

$X_t^{x,\varepsilon}$ is located near one of these points with overwhelming probability. The transitions between small neighborhoods of the equilibria are governed by the matrix

$$V_{ij} = \frac{1}{2} \inf \left\{ \int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2 ds, \quad \varphi_0 = O_i, \varphi_T = O_j \right\}, \quad i, j \in \{1, \dots, n\}, \quad (2)$$

where the infimum is taken over all $T \geq 0$ and all absolutely continuous functions $\varphi : [0, T] \rightarrow \mathbb{T}^d$. Loosely speaking, for each $i = 1, \dots, n$, $x \in D_i$, and $\lambda > 0$ (except a finite subset of λ 's, as discussed below) there is an index $k = k(x, \lambda) \in \{1, \dots, n\}$ such that

$$\text{dist}(X_{\exp(\lambda/\varepsilon)}^{x,\varepsilon}, O_k) \rightarrow 0 \quad \text{in probability as } \varepsilon \downarrow 0. \quad (3)$$

The equilibrium O_k is called the metastable state for the process $X_t^{x,\varepsilon}$ corresponding to the initial point x and time scale $\exp(\lambda/\varepsilon)$. The state O_k can be determined by comparing λ with certain linear expressions involving the numbers V_{ij} (see Chapter 6 of [4]). Typically, there is a finite set Λ of values of λ where transition from one metastable state to another happens, i.e., the notion of a metastable state is defined for $\lambda \in (0, \infty) \setminus \Lambda$.

On the other hand, for certain geometries of the unperturbed flow, it may happen that $V_{ij_1} = V_{ij_2}$ for some $j_1 \neq j_2$ or, more generally, the sums of two distinct collections of V_{ij} 's may be equal. The analysis of the asymptotics of $X_t^{x,\varepsilon}$ is then more intricate. The notion of rough symmetry, of which the simplest examples are due to geometric symmetries of the flow, was introduced and to some extent analyzed in [3]. In the presence of rough symmetry, metastable states may need to be replaced by metastable distributions between the asymptotically stable attractors.

One of the motivations of the current paper is to analyze an interesting situation where V_{ij} do not depend on j and, consequently, $X_t^{x,\varepsilon}$ is distributed between the neighborhoods of several equilibria at exponential time scales. Now this phenomenon is due not to geometric symmetries but to vanishing of the vector field in certain regions of the state space. Namely, let D_1, \dots, D_n be open connected domains with infinitely smooth boundaries ∂D_k , $k = 1, \dots, n$, on the d -dimensional torus \mathbb{T}^d . The closures \overline{D}_k are assumed to be disjoint. Let v be a vector field on \mathbb{T}^d that is equal to zero on $\mathbb{T}^d \setminus \bigcup_{k=1}^n \overline{D}_k$. It is assumed to be infinitely smooth in the sense that there is an infinitely smooth field on \mathbb{T}^d that agrees with v on $\bigcup_{k=1}^n \overline{D}_k$. Assume, for the moment, that all the points of D_k are attracted to an equilibrium $O_k \in D_k$, $k = 1, \dots, n$ (see Figure 1).

The quantities V_{ij} are easily seen not to depend on j since

$$\begin{aligned} \inf \left\{ \int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2 ds, \quad \varphi_0 \in \partial D_i, \varphi_T \in \partial D_j \right\} = \\ \inf \left\{ \int_0^T |\dot{\varphi}_s|^2 ds, \quad \varphi_0 \in \partial D_i, \varphi_T \in \partial D_j \right\} = 0, \end{aligned}$$

where the infimum is taken over all $T > 0$ and $\varphi \in C^1([0, T], \mathbb{T}^d)$. There are two issues involved in understanding the transitions of the process $X_t^{x,\varepsilon}$ between the neighborhoods

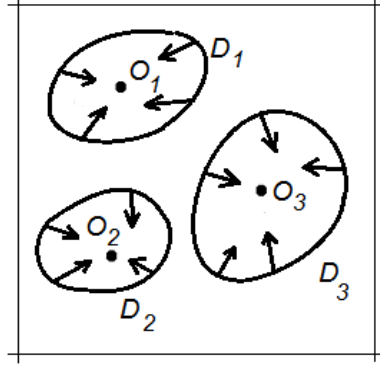


Figure 1: Torus with multiple trapping regions.

of different equilibria. The first issue is to describe how the process that starts near O_k exits the domain D_k . Fortunately, this question has been well-studied. One needs to look at the quasi-potential

$$V_k(x) = \frac{1}{2} \inf \left\{ \int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2 ds, \quad \varphi_0 = O_k, \varphi_T = x \right\}, \quad x \in \partial D_k, \quad (4)$$

where the infimum is taken over all $T \geq 0$ and all $\varphi \in C^1([0, T], \overline{D}_k)$. If the minimum of $V_k(x), x \in \partial D_k$, is achieved in a single point x_k , then the process starting near O_k exits D_k in a small neighborhood of x_k . The time to exit is of order $\exp(V_k(x_k)/\varepsilon)$. Even if the minimum is not achieved in a single point, there are cases when the exit from the domain is well understood (e.g., in the simplest example when v is spherically symmetric and D_k is a ball around O_k , also see [1] and references there). We'll simply assume that for each compact $K \subset D_k$, the exit time (appropriately re-scaled) and the exit location have limiting distributions that do not depend on the starting point within K , as in the case of a single minimum for the quasi-potential and in the symmetric case mentioned above.

The second issue concerns the transition between the domains D_k , i.e., the behavior of the process $X_t^{x, \varepsilon}$ on $\mathbb{T}^d \setminus \bigcup_{k=1}^n D_k$. In order to get a meaningful limiting object, we introduce a new process $Y_t^{x, \varepsilon}$ by running the clock only when $X_t^{x, \varepsilon} \in \mathbb{T}^d \setminus \bigcup_{k=1}^n D_k$. While this process obviously coincides with the Brownian motion away from boundary $\bigcup_{k=1}^n \partial D_k$, the regions D_k still play an important role by capturing the process when it reaches the boundary and then re-distributing it on ∂D_k . The description of the limiting behavior $Y_t^{x, \varepsilon}$ is the main result of this paper.

In Section 2 we describe the limiting process. It belongs to a peculiar class of processes that, it seems, has not been discussed in the literature previously. In Section 3 we prove the convergence of $Y_t^{x, \varepsilon}$ to the limit. In Section 4 we describe the asymptotics of the original process $X_t^{x, \varepsilon}$ at exponential time scales and make several additional remarks.

2 Description of the limiting process

In this section we define the family of processes X_t^x , which later will be proved to be the limiting processes for $Y_t^{x,\varepsilon}$ as $\varepsilon \downarrow 0$. Let $D_1, \dots, D_n \subset \mathbb{T}^d$ be open connected domains with infinitely smooth boundaries ∂D_k , $k = 1, \dots, n$. The closures \overline{D}_k are assumed to be disjoint. Let $U = \mathbb{T}^d \setminus \bigcup_{k=1}^n \overline{D}_k$. The closure of this domain will be denoted by \overline{U} . Let U' be the metric space obtained from \overline{U} by identifying all points of ∂D_k , making every ∂D_k , $k = 1, \dots, n$, into one point d_k .

The family of processes X_t^x , $x \in U'$, will be defined in terms of its generator. Since we expect X_t^x to coincide with a Brownian motion inside U , the generator coincides with $\frac{1}{2}\Delta$ on a certain class of functions. The domain, however, should be restricted by certain boundary conditions to account for non-trivial behavior of X_t^x on the boundary of U . We'll use the Hille-Yosida theorem stated here in the form that is convenient for considering closures of linear operators.

Theorem 2.1. *Let K be a compact space, $C(K)$ be the space of continuous functions on it. Suppose that a linear operator A on $C(K)$ has the following properties:*

- (a) *The domain $\mathcal{D}(A)$ is dense in $C(K)$;*
- (b) *The identity $\mathbf{1}$ belongs to $\mathcal{D}(A)$ and $A\mathbf{1} = 0$;*
- (c) *The maximum principle: If S is the set of points where a function $f \in \mathcal{D}(A)$ reaches its maximum, then $Af(x) \leq 0$ for at least one point $x \in S$.*
- (d) *For a dense set $\Psi \subseteq C(K)$ for every $\psi \in \Psi$ and every $\lambda > 0$ there exists a solution $f \in \mathcal{D}(A)$ of the equation $\lambda f - Af = \psi$.*

Then the operator A is closable and its closure \overline{A} is the infinitesimal generator of a unique semi-group of positivity-preserving operators T_t , $t \geq 0$, on $C(K)$ with $T_t\mathbf{1} = \mathbf{1}$, $\|T_t\| \leq 1$.

Suppose that we are given positive finite measures ν_1, \dots, ν_n concentrated on $\partial D_1, \dots, \partial D_n$, respectively. The Hille-Yosida theorem will be applied to the space $K = U'$. Let us define the linear operator A in $C(U')$. First we define its domain: it consists of all functions $f \in C(U')$ that satisfy the following conditions:

- (1) f is twice continuously differentiable in U ;
- (2) The limits of all the first and second order derivatives of f exist at all the points of the boundary $\partial U = \bigcup_{k=1}^n \partial D_k$;
- (3) There are constants g_1, \dots, g_n such that

$$\lim_{y \in U, \text{dist}(y, \partial D_k) \downarrow 0} \Delta f(y) = g_k, \quad k = 1, \dots, n;$$

- (4) For each $k = 1, \dots, n$,

$$\int_{\partial D_k} \langle \nabla f(x), n(x) \rangle \nu_k(dx) = 0, \tag{5}$$

where $n(x)$ is the unit exterior (with respect to U) normal at $x \in \partial D_k$.

For $f \in \mathcal{D}(A)$ and $x \in U'$, we define

$$Af = \begin{cases} \frac{1}{2}\Delta f(x), & \text{if } x \in U, \\ \frac{1}{2}g_k, & \text{if } x = d_k, \quad k = 1, \dots, n. \end{cases}$$

Let us check that the conditions of the Hille-Yosida theorem are satisfied.

(a) Consider the set G of functions g that are infinitely smooth and have the following property: for each $k = 1, \dots, n$ there is a set V_k open in U' such that $\partial D_k \subset V_k$ and g is constant on V_k . It is clear that $G \subset \mathcal{D}(A)$ and G is dense in $C(U')$.

(b) Clearly $\mathbf{1} \in \mathcal{D}(A)$ and $A\mathbf{1} = 0$.

(c) If f has a maximum at $x \in U$, it is clear that $\Delta f(x) \leq 0$. Now suppose that f has a maximum at d_k . We can view f as an element of $C^2(\bar{U})$ that is constant on each component of the boundary, in particular on ∂D_k . Note that $\langle \nabla f(x), n(x) \rangle$ is identically zero on ∂D_k , since otherwise it would be negative at some points due to (5), which would contradict the fact that f reaches its maximum on ∂D_k . Then the second derivative of f in the direction of n is non-positive at all points $x \in \partial D_k$. Since f is constant on ∂D_k , its second derivative in any direction tangential to the boundary is equal to zero. Therefore, $\Delta f(x) \leq 0$ for $x \in \partial D_k$, i.e., $Af(d_k) \leq 0$, as required.

(d) Let Ψ be the set of functions $\psi \in C(U')$ that have limits of all the first order derivatives as $y \in U, y \rightarrow x$, at all points $x \in \partial U$. It is clear that Ψ is dense in $C(U')$. Let $\tilde{f} \in C^2(\bar{U})$ be the solution of the equation $\lambda \tilde{f} - \frac{1}{2}\Delta \tilde{f} = \psi$ in U , $\tilde{f} = 0$ on ∂U . Let $h_k \in C^2(\bar{U})$ be the solution of the equation

$$\lambda h_k(x) - \frac{1}{2}\Delta h_k(x) = 0, \quad x \in U,$$

$$h_k(x) = 1, \quad x \in \partial D_k; \quad h_k(x) = 0, \quad x \in \partial U \setminus \partial D_k.$$

Let us look for the solution $f \in \mathcal{D}(A)$ of $\lambda f - Af = \psi$ in the form $f = \tilde{f} + \sum_{k=1}^n c_k h_k$. We get n linear equations for c_1, \dots, c_n . The solution is unique because of the maximum principle. Therefore, the determinant of the system is non-zero, and the solution exists for all the right hand sides.

Let \bar{A} be the closure of A . Let $T_t, t \geq 0$, be the corresponding semi-group on $C(U')$, whose existence is guaranteed by the Hille-Yosida theorem. By the Riesz-Markov-Kakutani representation theorem, for $x \in U'$ there is a measure $P(t, x, dy)$ on $(U', \mathcal{B}(U'))$ such that

$$(T_t f)(x) = \int_{U'} f(y) P(t, x, dy), \quad f \in C(U').$$

It is a probability measure since $T_t \mathbf{1} = \mathbf{1}$. Moreover, it can be easily verified that $P(t, x, B)$ is a Markov transition function. Let $X_t^x, x \in U'$, be the corresponding Markov family. In order to show that a modification with continuous trajectories exists, it is enough to check that $\lim_{t \downarrow 0} P(t, x, B)/t = 0$ for each closed set B that doesn't contain x (Theorem

I.5 of [5], see also [2]). Let $f \in \mathcal{D}(A)$ be a non-negative function that is equal to one on B and whose support doesn't contain x . Then

$$\lim_{t \downarrow 0} \frac{P(t, x, B)}{t} \leq \lim_{t \downarrow 0} \frac{(T_t f)(x) - f(x)}{t} = Af(x) = 0,$$

as required. Thus X_t^x can be assumed to have continuous trajectories.

3 Convergence of the trace of the process

In this section we prove the convergence of $Y_t^{x,\varepsilon}$, obtained from $X_t^{x,\varepsilon}$ by running the clock only when the process is in \overline{U} , to the limiting process X_t^x .

Let v be a vector field that is smooth in $\bigcup_{k=1}^n \overline{D}_k$ (i.e., it admits a smooth continuation from $\bigcup_{k=1}^n \overline{D}_k$ to the whole space) and is equal to zero outside $\bigcup_{k=1}^n \overline{D}_k$.

Let

$$a(x) = \langle v(x), n(x) \rangle, \quad x \in \partial U,$$

where n is the unit exterior (with respect to U) normal to the boundary. We'll assume that $a(x) > 0$ for all $x \in \partial U$. Recall that the process $X_t^{x,\varepsilon}$ is defined via

$$dX_t^{x,\varepsilon} = \frac{1}{\varepsilon} v(X_t^{x,\varepsilon}) dt + dW_t, \quad X_0^{x,\varepsilon} = x.$$

For $B \subset \mathbb{R}$, let $\tau^{x,\varepsilon}(B) = \inf\{t \geq 0 : X_t^{x,\varepsilon} \in B\}$. Let $\mu_k^{x,\varepsilon}$ be the measure on ∂D_k induced by $X_{\tau^{x,\varepsilon}(\partial D_k)}^{x,\varepsilon}$. We'll assume that there are measures μ_k , $k = 1, \dots, n$, such that for each compact set $K \subset D_k$ and each continuous function φ on ∂D_k we have

$$\lim_{\varepsilon \downarrow 0} \int_{\partial D_k} \varphi d\mu_k^{x,\varepsilon} = \int_{\partial D_k} \varphi d\mu_k \quad (6)$$

uniformly in $x \in K$. Define

$$s(t) = \inf(s : \lambda(u : u \leq s, X_u^{x,\varepsilon} \in \overline{U}) > t),$$

where λ is the Lebesgue measure on the real line, and

$$Y_t^{x,\varepsilon} = X_{s(t)}^{x,\varepsilon}. \quad (7)$$

Thus $Y_t^{x,\varepsilon}$ is a right-continuous process with values in \overline{U} , which also can be viewed as a right-continuous U' -valued process. It can be obtained from $X_t^{x,\varepsilon}$ by running the clock only when $X_t^{x,\varepsilon}$ is in \overline{U} . Define the measures ν_k via

$$\nu_k(dx) = (2a(x))^{-1} \mu_k(dx), \quad x \in \partial D_k. \quad (8)$$

Let X_t^x be the Markov family of continuous U' -valued processes defined above, corresponding to the measures ν_k , $k = 1, \dots, n$. The main result of this section is the following.

Theorem 3.1. *For each $x \in U'$, the measures on $C([0, \infty), U')$ induced by the processes $Y_t^{x, \varepsilon}$ converge weakly, as $\varepsilon \downarrow 0$, to the measure induced by X_t^x .*

The key ingredient in the proof of this theorem is the following lemma.

Lemma 3.2. *Suppose that $f \in \mathcal{D}(A)$. Then*

$$\lim_{\varepsilon \downarrow 0} \mathbb{E} \left(f(Y_t^{x, \varepsilon}) - f(x) - \frac{1}{2} \int_0^t \Delta f(Y_u^{x, \varepsilon}) du \right) = 0 \quad (9)$$

for each $t \geq 0$, uniformly in $x \in \bar{U}$.

Proof. For the sake of notational simplicity, we'll assume that there is just one domain where the vector field v is non-zero. This does not lead to any loss of generality as the proof in the case of multiple domains is similar. We'll denote the domain by D and will drop the subscript k from the notations everywhere. For example, (5) now takes the form

$$\int_{\partial D} \langle \nabla f(x), n(x) \rangle d\nu(x) = 0. \quad (10)$$

Let $S_r = \{x \in \bar{U} : \text{dist}(x, \partial D) = r\}$ for $r \geq 0$, $S_r = \{x \in D : \text{dist}(x, \partial D) = -r\}$ for $r < 0$. These are smooth surfaces if r is sufficiently small. Let $\Gamma_r = \{x \in \bar{U} : \text{dist}(x, \partial D) \leq r\}$ for $r \geq 0$.

Let $\sigma_0^{x, \varepsilon} = 0$, $\tau_1^{x, \varepsilon} = \tau^{x, \varepsilon}(\partial D)$, $\sigma_n^{x, \varepsilon} = \inf(t \geq \tau_n^{x, \varepsilon} : X_t^{x, \varepsilon} \in S_{\sqrt{\varepsilon}})$, $n \geq 1$, while $\tau_n^{x, \varepsilon} = \inf(t \geq \sigma_{n-1}^{x, \varepsilon} : X_t^{x, \varepsilon} \in \partial D)$, $n \geq 2$. Then

$$\begin{aligned} & \mathbb{E} \left(f(Y_t^{x, \varepsilon}) - f(x) - \frac{1}{2} \int_0^t \Delta f(Y_u^{x, \varepsilon}) du \right) = \\ & \mathbb{E} \sum_{n=1}^{\infty} \left(f(X_{\tau_n^{x, \varepsilon} \wedge s(t)}^{x, \varepsilon}) - f(X_{\sigma_{n-1}^{x, \varepsilon} \wedge s(t)}^{x, \varepsilon}) - \frac{1}{2} \int_{\sigma_{n-1}^{x, \varepsilon} \wedge s(t)}^{\tau_n^{x, \varepsilon} \wedge s(t)} \Delta f(X_u^{x, \varepsilon}) du \right) + \\ & \mathbb{E} \sum_{n=1}^{\infty} \left(f(X_{\sigma_n^{x, \varepsilon} \wedge s(t)}^{x, \varepsilon}) - f(X_{\tau_n^{x, \varepsilon} \wedge s(t)}^{x, \varepsilon}) - \frac{1}{2} \int_{\tau_n^{x, \varepsilon} \wedge s(t)}^{\sigma_n^{x, \varepsilon} \wedge s(t)} \Delta f(X_u^{x, \varepsilon}) du \right), \end{aligned} \quad (11)$$

where we put $\Delta f \equiv 0$ on D . The first expectation on the right hand side is equal to zero since $X_t^{x, \varepsilon}$ is pure Brownian motion on \bar{U} . Next, we prove several statements primarily concerning the stopping times $\sigma_n^{x, \varepsilon}$ and the behavior of the process near ∂D .

1) There is $c = c(t) > 0$ such that

$$\mathbb{P}(\sigma_n^{x, \varepsilon} \leq s(t)) \leq \exp(-cn\sqrt{\varepsilon}), \quad x \in \bar{U}, \quad n \geq 2. \quad (12)$$

Indeed, since ∂D is smooth, there is $r > 0$ such that the ball of radius r tangent to ∂D at x lies entirely in \bar{U} . Let η^ε be the time it takes a Brownian motion starting inside a

ball of radius r at a distance $\sqrt{\varepsilon}$ from the boundary to reach the boundary. It is easy to see that there is $c = c(t) > 0$ such that

$$P(\eta^\varepsilon \leq t) \leq \exp(-c\sqrt{\varepsilon}).$$

Therefore, if η_k^ε , $k \geq 1$, is a sequence of such independent random variables, then

$$P(\eta_1^\varepsilon + \dots + \eta_n^\varepsilon \leq t) \leq \exp(-cn\sqrt{\varepsilon}), \quad n \geq 1. \quad (13)$$

By our construction, $P(\tau_n^{x,\varepsilon} - \sigma_{n-1}^{x,\varepsilon} > z | X_{\sigma_{n-1}^{x,\varepsilon}}^{x,\varepsilon}) \geq P(\eta^\varepsilon > z)$ for each $n \geq 2$ and $z > 0$. Therefore, estimate (12), with $\tau_{n+1}^{x,\varepsilon}$ instead of $\sigma_n^{x,\varepsilon}$, follows from (13) and the strong Markov property. Thus, the original formula (12) also holds, with a different constant c .

2) Consider an auxiliary one-dimensional process Z_t^z that satisfies

$$dZ_t^z = dB_t - v\chi_{(-\infty,0)}(Z_t^z)dt, \quad Z_0^z = z,$$

where $v > 0$ and B_t is a one-dimensional Brownian motion. Also consider its perturbation defined via

$$d\tilde{Z}_t^z = dB_t - v\chi_{(-\infty,0)}(\tilde{Z}_t^z)dt + A_t^z dt, \quad \tilde{Z}_0^z = z, \quad (14)$$

where A_t^z is a continuous adapted process satisfying $|A_t^z| \leq v/2$. It is easy to see that for each $\eta > 0$ there is $z_0 < 0$ such that

$$P(\sup_{t \geq 0} \tilde{Z}_t^{z_0} \geq 0) \leq \eta, \quad (15)$$

while for each $z_0 < 0$ there is t_0 such that for $z \in [z_0, 1]$,

$$P(\tilde{Z}_t^z \text{ reaches } \{z_0\} \cup \{1\} \text{ before time } t_0) \geq 1 - \eta. \quad (16)$$

A direct calculation shows that

$$P(\sup_{t \geq 0} Z_t^0 \geq 1) = (1 + 2v)^{-1}.$$

Fix such z_0 that (15) holds and

$$P(Z_t^0 \text{ reaches } 1 \text{ before reaching } z_0) \geq (1 + 2v)^{-1} - \eta. \quad (17)$$

By the Girsanov formula, whose use is justified by (16) with \tilde{Z}_t^z replaced by Z_t^0 , there is $\varkappa > 0$ such that

$$P(\tilde{Z}_t^0 \text{ reaches } 1 \text{ before reaching } z_0) \in [(1 + 2v)^{-1} - 3\eta, (1 + 2v)^{-1} + 3\eta], \quad (18)$$

provided that $|A_t^0| \leq \varkappa$. Below we'll encounter a related ε -dependent process. Namely, suppose that $(\tilde{Z}_t^{z,\varepsilon}, \hat{\tilde{Z}}_t^{z,\varepsilon}) \in \mathbb{R} \times \mathbb{R}^{d-1}$ satisfy

$$d\tilde{Z}_t^{z,\varepsilon} = d\tilde{B}_t - \frac{v}{\varepsilon}\chi_{(-\infty,0)}(\tilde{Z}_t^{z,\varepsilon})dt + \frac{\tilde{A}_t^z}{\varepsilon}dt, \quad \tilde{Z}_0^{z,\varepsilon} = \tilde{z}, \quad (19)$$

$$d\widehat{Z}_t^{z,\varepsilon} = \sigma(\widetilde{Z}_t^{z,\varepsilon}, \widehat{Z}_t^{z,\varepsilon})d\widehat{B}_t + \frac{\widehat{A}_t^z}{\varepsilon}dt, \quad \widehat{Z}_0^{z,\varepsilon} = \widehat{z}, \quad (20)$$

where \widetilde{B}_t is a one-dimensional Brownian motion, \widehat{B}_t is d -dimensional Brownian motion, possibly correlated with \widetilde{B}_t , σ is a $(d-1) \times d$ matrix, and $z = (\widetilde{z}, \widehat{z}) \in \mathbb{R}^d$. We'll assume that there is $C > 0$ such that

$$|\widetilde{A}_t^z| \leq v/2, \quad |\sigma|, |\widehat{A}_t^z| \leq C. \quad (21)$$

For $A \subset \mathbb{R}$, let $\widetilde{\tau}^{z,\varepsilon}(A) = \inf\{t \geq 0 : \widetilde{Z}_t^{z,\varepsilon} \in A\}$. We claim that there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E} \left(\lambda(t : \widetilde{Z}_t^{0,\varepsilon} \in [0, \varepsilon], t \leq \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\} \cup \{\varepsilon\})) \right) \leq L\varepsilon^2, \quad (22)$$

provided that $\varepsilon \leq \varepsilon_0$. If, additionally,

$$|\widetilde{A}_t^0| \leq \varkappa(\varepsilon) \quad \text{whenever} \quad |\widetilde{Z}_t^{0,\varepsilon}| + |\widehat{Z}_t^{0,\varepsilon}| \leq \sqrt{\varepsilon}, \quad \text{with } \varkappa(\varepsilon) \downarrow 0, \text{ as } \varepsilon \downarrow 0, \quad (23)$$

then

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\})) = (1 + 2v)^{-1}. \quad (24)$$

Indeed, let $\eta \in (0, 1)$. Find $z_0 < 0$ such that (17) holds and

$$\mathbb{P}(\widetilde{\tau}^{z_0\varepsilon,\varepsilon}(\{0\}) < \infty) \leq \eta, \quad (25)$$

which is possible by (15) and the scaling-invariance of the Brownian motion. By (16), we can find t_0 such that for $z \in [z_0, 1]$,

$$\mathbb{P}(\widetilde{\tau}^{z\varepsilon,\varepsilon}(\{z_0\varepsilon\} \cup \{\varepsilon\}) \leq t_0\varepsilon^2) \geq 1 - \eta. \quad (26)$$

The combination of these two inequalities and the strong Markov property (whose use is allowed since a time-shift of the process \widetilde{A}_t^z is also bounded by $v/2$) imply (22). By (23), (26), (19), and (20),

$$\mathbb{P}(\sup_{t \leq \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\} \cup \{\varepsilon\})} |\widetilde{A}_t^0| \leq \varkappa(\varepsilon)) \geq \mathbb{P}(\sup_{t \leq \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\} \cup \{\varepsilon\})} (|\widetilde{Z}_t^{0,\varepsilon}| + |\widehat{Z}_t^{0,\varepsilon}|) \leq \sqrt{\varepsilon}) \geq 1 - 2\eta$$

for all sufficiently small ε . Therefore, by (18), which can be applied after re-scaling,

$$\mathbb{P}(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\})) \in [(1 + 2v)^{-1} - 5\eta, (1 + 2v)^{-1} + 5\eta].$$

Together with (25), this implies that

$$\mathbb{P}(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\})) \in [(1 + 2v)^{-1} - 6\eta, (1 + 2v)^{-1} + 6\eta].$$

Since η was arbitrary, this implies (24).

From the presence of a strong drift to the left in (19), it easily follows that there is $L > 0$ such that

$$\sup_{z \in [-\sqrt{\varepsilon}, \varepsilon]} \mathbb{E} \tilde{\tau}^{z, \varepsilon}(\{-\sqrt{\varepsilon}\} \cup \{\varepsilon\}) \leq L\varepsilon^{\frac{3}{2}}. \quad (27)$$

From (27), (19), and (20) it follows that

$$\lim_{\varepsilon \downarrow 0} \mathbb{P} \left(\sup_{t \leq \tilde{\tau}^{0, \varepsilon}(\{-\sqrt{\varepsilon}\} \cup \{\varepsilon\})} (|\tilde{Z}_t^{0, \varepsilon}| + |\hat{Z}_t^{0, \varepsilon}|) > \varepsilon^{\frac{1}{3}} \right) = 0. \quad (28)$$

3) Let us show that for each sufficiently small $\delta > 0$ there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E} (\lambda(t : X_t^{x, \varepsilon} \in \bar{U}, t \leq \min(\tau^{x, \varepsilon}(S_{-\delta}), \tau^{x, \varepsilon}(S_\varepsilon))) \leq L\varepsilon^2, \quad x \in \partial D, \quad (29)$$

provided that $\varepsilon \leq \varepsilon_0$, and that

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(\tau^{x, \varepsilon}(S_\varepsilon) < \tau^{x, \varepsilon}(S_{-\delta})) = (1 + 2a(x))^{-1} \quad \text{uniformly in } x \in \partial D. \quad (30)$$

First observe that

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-2} \mathbb{P}(\tau^{x, \varepsilon}(S_\varepsilon) < \tau^{x, \varepsilon}(S_{-\delta})) = 0 \quad \text{uniformly in } x \in S_{-\sqrt{\varepsilon}}, \quad (31)$$

due to the presence of the strong drift inside D . Also, there is $L > 0$ such that

$$\mathbb{E} \min(\tau^{x, \varepsilon}(\partial D), \tau^{x, \varepsilon}(S_\varepsilon)) \leq L\varepsilon^2, \quad x \in \Gamma_\varepsilon, \quad (32)$$

for all sufficiently small ε , since the process is pure Brownian motion on U .

Let us describe a change of coordinates in a neighborhood of a point $x \in \partial D$. Let $V_\varepsilon^r = [-\sqrt{\varepsilon}, \varepsilon] \times B_r \subset \mathbb{R}^d$, where $B_r \subset \mathbb{R}^{d-1}$ is the closed ball of radius r centered at the origin. Let m_x be an isometric mapping of B_r to the $(d-1)$ -dimensional ball of radius r centered at x in the tangent plane to ∂D at x . For $y \in B_r$, we take the straight line passing through $m_x(y)$ and the point on ∂D closest to $m_x(y)$ (this line is perpendicular to ∂D if $m_x(y) \notin \partial D$; we define it as the perpendicular if $m_x(y) \in \partial D$). For $z \in [-\sqrt{\varepsilon}, \varepsilon]$ and $y \in B_r$, define $\varphi_x(z, y)$ as the point on the perpendicular that belongs to S_z . If r and ε are sufficiently small, then φ_x is a diffeomorphism from V_ε^r to a domain $U_\varepsilon^r(x)$ for each x .

For $z \in V_\varepsilon^r$, let $\bar{X}_t^{z, \varepsilon} = \varphi^{-1}(X_t^{\varphi(z), \varepsilon})$ be the process written in the new coordinates (stopped when it reaches the boundary of V_ε^r). It satisfies

$$d\bar{X}_t^{z, \varepsilon} = \frac{1}{\varepsilon} \bar{v}(\bar{X}_t^{z, \varepsilon}) \chi_{[-\sqrt{\varepsilon}, 0] \times \mathbb{R}^{d-1}}(\bar{X}_t^{z, \varepsilon}) dt + \bar{\beta}(\bar{X}_t^{z, \varepsilon}) dt + \bar{\sigma}(\bar{X}_t^{z, \varepsilon}) d\bar{W}_t, \quad \bar{X}_0^{z, \varepsilon} = z.$$

The coefficients $\bar{v}, \bar{\beta}, \bar{\sigma}$ are bounded in $C^1(V_\varepsilon^r)$ (the change of coordinates, and therefore the coefficients, depend on x , but the bound is uniform in x) and satisfy: $\bar{v}_1(0) = -a(x)$, $\bar{\sigma}(0)$ is an orthogonal matrix. Let

$$\alpha(z) = (\sigma_{11}^2(z) + \dots + \sigma_{1d}^2(z))^{-\frac{1}{2}}.$$

Note that this is a smooth function such that $\alpha(0) = 1$. The process

$$d\tilde{X}_t^{z,\varepsilon} = \frac{1}{\varepsilon} \alpha^2 \bar{v} \chi_{[-\sqrt{\varepsilon}, 0] \times \mathbb{R}^{d-1}}(\tilde{X}_t^{z,\varepsilon}) dt + \alpha^2 \bar{\beta}(\tilde{X}_t^{z,\varepsilon}) dt + \alpha \bar{\sigma}(\tilde{X}_t^{z,\varepsilon}) d\bar{W}_t, \quad \tilde{X}_0^{z,\varepsilon} = z,$$

is different from $\bar{X}_t^{z,\varepsilon}$ by a random change of time. The coefficients of $\tilde{X}_t^{z,\varepsilon}$ can be extended from $V_\varepsilon^r = [-\sqrt{\varepsilon}, \varepsilon] \times B_r \subset \mathbb{R}^d$ to $[-\sqrt{\varepsilon}, \varepsilon] \times \mathbb{R}^{d-1}$ as continuous functions by requiring that they don't vary in the radial direction outside B_r . This way the process $\tilde{X}_t^{z,\varepsilon}$ can be defined until the time it reaches the boundary of $[-\sqrt{\varepsilon}, \varepsilon] \times \mathbb{R}^{d-1}$. Let $\tilde{Z}_t^{z,\varepsilon}$ be the first coordinate of $\tilde{X}_t^{z,\varepsilon}$ and $\hat{Z}_t^{z,\varepsilon}$ be the vector consisting of the remaining $d-1$ coordinates of $\tilde{X}_t^{z,\varepsilon}$. Note that the process $(\tilde{Z}_t^{z,\varepsilon}, \hat{Z}_t^{z,\varepsilon})$ can be written in the form (19)-(20) with the coefficients satisfying (21), (23), provided that r and ε chosen to be sufficiently small, independently of x .

By (28),

$$\lim_{\varepsilon \downarrow 0} \mathbb{P} \left(\min(\tau^{x,\varepsilon}(S_{-\sqrt{\varepsilon}}), \tau^{x,\varepsilon}(S_\varepsilon)) > \tau^{x,\varepsilon}(U_\varepsilon^r(x)) \right) = 0, \quad (33)$$

and it is not difficult to see that the limit is uniform in $x \in \partial D$. Since α is bounded from above and below, from (22) and (24) it follows that there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E} \left(\lambda(t : X_t^{x,\varepsilon} \in \bar{U}, t \leq \tau^{x,\varepsilon}(U_\varepsilon^r(x))) \right) \leq L\varepsilon^2, \quad x \in \partial D, \quad (34)$$

provided that $\varepsilon \leq \varepsilon_0$, and that

$$\lim_{\varepsilon \downarrow 0} \mathbb{P} \left(\tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\sqrt{\varepsilon}}) \right) = (1 + 2a(x))^{-1}, \quad x \in \partial D. \quad (35)$$

The convergence is uniform in x since the dependence of the process $(\tilde{Z}_t^{z,\varepsilon}, \hat{Z}_t^{z,\varepsilon})$ on x manifests itself through the value of $v = a(x)$ and through the values of $C, \varkappa(\varepsilon)$ in (21), (23) in a way that doesn't affect the applicability of (22) and (24).

The strong Markov property of the process $X_t^{x,\varepsilon}$, together with (31), (32), and (33), allows us to obtain (29) from (34) and (30) from (35).

4) Let us get a bound on $\mathbb{E}\lambda(u : u \leq \sigma_1^{x,\varepsilon}, X_u^{x,\varepsilon} \in \bar{U}), x \in \partial D$. Since the process is pure Brownian motion on U , there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E} \min(\tau^{x,\varepsilon}(\partial D), \tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})) \leq L\varepsilon^{\frac{3}{2}}, \quad x \in S_\varepsilon, \quad (36)$$

provided that $\varepsilon \leq \varepsilon_0$, while

$$\lim_{\varepsilon \downarrow 0} (\varepsilon^{-\frac{1}{2}} \mathbb{P}(\tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}}) < \tau^{x,\varepsilon}(\partial D))) = 1, \quad \text{uniformly in } x \in S_\varepsilon. \quad (37)$$

Let δ be sufficiently small for (29) and (30) to hold. Formulas (36), (37) together with (29), (30), and the strong Markov property of the process, imply that there are $\varepsilon_0 > 0$ and $L > 0$ such that for all sufficiently small ε ,

$$\mathbb{E}\lambda(u : u \leq \sigma_1^{x,\varepsilon}, X_u^{x,\varepsilon} \in \bar{U}) \leq L\varepsilon, \quad x \in \partial D. \quad (38)$$

Let $\xi_n^{x,\varepsilon} = \lambda(u : \tau_n^{x,\varepsilon} \leq u \leq \sigma_n^{x,\varepsilon}, X_u^{x,\varepsilon} \in \overline{U})$. Formulas (38) and (12), together with the strong Markov property of the process, imply that there is $c = c(t)$ such that

$$\mathbb{E} \sum_{n=0}^{\infty} \xi_{n+1}^{x,\varepsilon} \chi_{\{\sigma_n^{x,\varepsilon} \leq s(t)\}} \leq c\sqrt{\varepsilon}, \quad x \in \overline{U}. \quad (39)$$

5) We claim that for sufficiently small δ there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E} \left(|X_{\tau^{x,\varepsilon}(S_\varepsilon)}^{x,\varepsilon} - x|^2, \tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\delta}) \right) \leq L\varepsilon^2, \quad x \in \partial D, \quad (40)$$

provided that $\varepsilon \leq \varepsilon_0$. Let us sketch a proof of this statement. First, by observing the process in the ε -neighborhood of ∂D , it is easy to see that

$$\mathbb{E} |X_{\tau^{x,\varepsilon}(S_\varepsilon) \wedge \tau^{x,\varepsilon}(S_{-\varepsilon})}^{x,\varepsilon} - x|^2 \leq L\varepsilon^2, \quad x \in \partial D. \quad (41)$$

From the the presence of the strong drift inside D , it follows that there is $c > 0$ such that

$$\mathbb{P} \left(\tau^{x,\varepsilon}(\partial D) > t\varepsilon^2, \tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta}) \right) \leq e^{-ct}, \quad x \in \partial S_{-\varepsilon}, \quad t \geq 0, \quad (42)$$

and, consequently,

$$\mathbb{E} \left(|X_{\tau^{x,\varepsilon}(\partial D)}^{x,\varepsilon} - x|^2, \tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta}) \right) \leq L\varepsilon^2, \quad x \in \partial S_{-\varepsilon}. \quad (43)$$

Also note that there is $c > 0$ such that

$$\mathbb{P} \left(\tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta}) \right) \leq 1 - c, \quad x \in \partial S_{-\varepsilon}. \quad (44)$$

By considering consecutive visits of the process to $S_{-\varepsilon}$ and ∂D , employing (41), (43), (44), and using the strong Markov property, we obtain (40).

Take the compact set $K \subset D$ such that $\partial K = S_{-\delta}$, where δ is sufficiently small for (29), (30), and (40) to hold. Observe that, since the process is pure Brownian motion in U , by (36),

$$\mathbb{E} |X_{\tau^{x,\varepsilon}(\partial D) \wedge \tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})}^{x,\varepsilon} - x|^2 = d \cdot \mathbb{E}(\tau^{x,\varepsilon}(\partial D) \wedge \tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})) \leq L\varepsilon^{\frac{3}{2}}, \quad x \in S_\varepsilon. \quad (45)$$

Therefore, by (40), (30), and (37), it follows from the strong Markov property that

$$\lim_{\varepsilon \downarrow 0} (\varepsilon^{-\frac{1}{2}} \mathbb{P}(\sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K))) = \sum_{n=1}^{\infty} \left(\frac{1}{1 + 2a(x)} \right)^n = \frac{1}{2a(x)}, \quad \text{uniformly in } x \in \partial D. \quad (46)$$

6) Combining (30), (37), (40), and (45), and using the strong Markov property, we obtain that there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathbb{E}(|X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - x|^2, \sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K)) \leq L\varepsilon^{\frac{3}{2}}, \quad x \in \partial D, \quad (47)$$

provided that $\varepsilon \leq \varepsilon_0$. Therefore,

$$\mathbb{E}(|X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - x|, \sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K)) \leq L\varepsilon, \quad x \in \partial D \quad (48)$$

provided that $\varepsilon \leq \varepsilon_0$, with a different constant L .

7) Let \bar{f} be the value of f on ∂D . Let us show that

$$\lim_{\varepsilon \downarrow 0} (\varepsilon^{-\frac{1}{2}} \mathbb{E}(f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f})) = 0, \quad \text{uniformly in } x \in \partial D. \quad (49)$$

Introduce the following two sequences of stopping times: $\bar{\tau}_1^{x,\varepsilon} = \tau^{x,\varepsilon}(\partial D)$, $\bar{\sigma}_n^{x,\varepsilon} = \inf(t \geq \bar{\tau}_n^{x,\varepsilon} : X_t^{x,\varepsilon} \in K)$, $n \geq 1$, while $\bar{\tau}_n^{x,\varepsilon} = \inf(t \geq \bar{\sigma}_{n-1}^{x,\varepsilon} : X_t^{x,\varepsilon} \in \partial D)$, $n \geq 2$. Then

$$\begin{aligned} \mathbb{E}f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f} &= \sum_{n=1}^{\infty} \mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \bar{\tau}_n^{x,\varepsilon} < \sigma_1^{x,\varepsilon} < \bar{\tau}_{n+1}^{x,\varepsilon}) = \\ &= \mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon}) + \sum_{n=2}^{\infty} \mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \sigma_1^{x,\varepsilon} < \bar{\tau}_{n+1}^{x,\varepsilon} | \sigma_1^{x,\varepsilon} > \bar{\sigma}_{n-1}^{x,\varepsilon}) \mathbb{P}(\sigma_1^{x,\varepsilon} > \bar{\sigma}_{n-1}^{x,\varepsilon}). \end{aligned}$$

By (46) and the strong Markov property of the process, there is $c > 0$ such that

$$\mathbb{P}(\sigma_1^{x,\varepsilon} > \bar{\sigma}_{n-1}^{x,\varepsilon}) \leq (1 - c\sqrt{\varepsilon})^{n-1}.$$

Also observe that

$$\begin{aligned} |\mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \sigma_1^{x,\varepsilon} < \bar{\tau}_{n+1}^{x,\varepsilon} | \sigma_1^{x,\varepsilon} > \bar{\sigma}_{n-1}^{x,\varepsilon})| &\leq \sup_{x \in K} |\mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon})| \leq \\ \sup_{x \in K} |\mathbb{E}(\langle \nabla f(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}), X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon} \rangle, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon})| &+ C \sup_{x \in \partial D} \mathbb{E}(|X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - x|^2, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon}), \end{aligned}$$

where C depends on the $C^2(\bar{U})$ -norm of f . The second term on the right hand side is bounded by $C\varepsilon^{\frac{3}{2}}$, with a different constant C , using (47). In order to estimate the first term, we notice that if $x \in \partial D$ and $y \in S_{\sqrt{\varepsilon}}$, then, since $\nabla f(x)$ is orthogonal to ∂D ,

$$|\langle \nabla f(x), y - x \rangle + \sqrt{\varepsilon} \langle \nabla f(x), n(x) \rangle| \leq c|x - y|^2$$

for some $c > 0$, where $n(x)$ is the unit inward (with respect to D) normal to the boundary at x . Therefore, for $x \in K$,

$$|\mathbb{E}(\langle \nabla f(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}), X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon} \rangle, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon}) + \sqrt{\varepsilon} \mathbb{E}(\langle \nabla f(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}), n(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}) \rangle, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon})| \leq$$

$$CE(|X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon} - X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}|^2, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon}),$$

where C depends on the $C^1(\bar{U})$ -norm of f . The right hand side is bounded by $C\varepsilon^{\frac{3}{2}}$, with a different constant C , using (47) and the strong Markov property. Observe that

$$\limsup_{\varepsilon \downarrow 0} |\varepsilon^{-\frac{1}{2}} \mathbb{E}(\langle \nabla f(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}), n(X_{\bar{\tau}_1^{x,\varepsilon}}^{x,\varepsilon}) \rangle, \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon})| = 0$$

by (6), (10), (46), and the strong Markov property of the process. By (48), we also have

$$\limsup_{\varepsilon \downarrow 0} |\varepsilon^{-\frac{1}{2}} \mathbb{E}((f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}), \sigma_1^{x,\varepsilon} < \bar{\tau}_2^{x,\varepsilon})| = 0.$$

Combining the estimates above, we obtain (49).

Let us examine the second term in the right hand side of (11). From (39) and the boundedness of Δf it follows that

$$\lim_{\varepsilon \downarrow 0} \mathbb{E} \sum_{n=1}^{\infty} \int_{\tau_n^{x,\varepsilon} \wedge s(t)}^{\sigma_n^{x,\varepsilon} \wedge s(t)} \Delta f(X_u^{x,\varepsilon}) du = 0.$$

It remains to show that

$$\lim_{\varepsilon \downarrow 0} \mathbb{E} \sum_{n=1}^{\infty} \left(f(X_{\sigma_n^{x,\varepsilon} \wedge s(t)}^{x,\varepsilon}) - f(X_{\tau_n^{x,\varepsilon} \wedge s(t)}^{x,\varepsilon}) \right) = 0. \quad (50)$$

Introduce the stopping time

$$s'(t) = \begin{cases} \sigma_n^{x,\varepsilon} & \text{if } \tau_n^{x,\varepsilon} < s(t) \leq \sigma_n^{x,\varepsilon} \\ s(t) & \text{otherwise.} \end{cases}$$

Now (50) will follow if we show that

$$\lim_{\varepsilon \downarrow 0} \mathbb{E} \sum_{n=1}^{\infty} \left(f(X_{\sigma_n^{x,\varepsilon} \wedge s'(t)}^{x,\varepsilon}) - \bar{f} \right) = 0, \quad (51)$$

since the difference between (51) and (50) is estimated from above by $2 \sup_{x \in S_{\sqrt{\varepsilon}}} |f(x) - \bar{f}|$, which goes to zero as $\varepsilon \downarrow 0$. Let $N^{x,\varepsilon} = \max(n : \sigma_n^{x,\varepsilon} \leq s'(t))$. By the strong Markov property,

$$\sup_{x \in \bar{U}} \mathbb{E} \sum_{n=1}^{\infty} \left(f(X_{\sigma_n^{x,\varepsilon} \wedge s'(t)}^{x,\varepsilon}) - \bar{f} \right) \leq \sup_{x \in \bar{U}} \mathbb{E} N^{x,\varepsilon} \sup_{x \in \partial D} \mathbb{E} (f(X_{\sigma_1^{x,\varepsilon}}^{x,\varepsilon}) - \bar{f}).$$

The right hand side tends to zero by (12) and (49). \square

Proof of Theorem 3.1. Recall that Ψ is the set of functions $\psi \in C(U')$ that have limits of all the first order derivatives as $y \in U, y \rightarrow x$, at all points $x \in \partial U$. This is a measure-defining class of functions on U' , i.e., if μ_1 and μ_2 satisfy $\int_{U'} \psi d\mu_1 = \int_{U'} \psi d\mu_2$ for every $\psi \in \Psi$, then $\mu_1 = \mu_2$. As shown in Section 2, for every $\psi \in \Psi$ and every $\lambda > 0$, there is $f \in \mathcal{D}(A)$ that satisfies $\lambda f - Af = \psi$. We have demonstrated that (9) holds for $f \in \mathcal{D}(A)$. By Lemma 3.1. in Chapter 8 of [4], this is sufficient to guarantee the convergence if, in addition, the family $\{Y_t^{x,\varepsilon}\}, \varepsilon > 0, x \in U'$ is tight. The tightness, however, is clear since the processes coincide with Brownian motion inside U , while all the points of ∂U are identified. \square

4 Applications, generalizations, and remarks

4.1 The behavior of the process at exponential time scales

Let us now discuss the behavior, as $\varepsilon \downarrow 0$, of the original process $X_t^{x,\varepsilon}$ (rather than its trace $Y_t^{x,\varepsilon}$ on \overline{U}). If the process starts in a small neighborhood of O_k , then it takes time of order $\exp(V_k/\varepsilon)$ (in the sense of logarithmic equivalence) for it to reach ∂D_k , where $V_k = \inf_{x \in \partial D_k} V_k(x)$ and $V_k(x)$ is the quasi-potential defined in (4). Thus it is reasonable to study the behavior of $X_t^{x,\varepsilon}$ at exponential time scales, i.e., at times of order $\exp(\lambda/\varepsilon)$ with fixed λ .

The transitions between small neighborhoods of the equilibria are governed by the matrix V_{ij} defined in (2). In our case, $V_{ij} = V_i$ for all i, j , as explained in the Introduction. Because of this “rough symmetry” ([3]), the notion of a metastable state (see (3)) should be replaced by that of a metastable distribution between the equilibria.

To describe the metastable distribution for a given initial point $x \in \mathbb{T}^d$ and time scale $\exp(\lambda/\varepsilon)$ with $\lambda > 0$, assume that $V_1 < V_2 < \dots < V_n$, and put $V_0 = 0, V_{n+1} = \infty$. We introduce the following non-standard boundary problem, which will be referred to as the (k, j) -problem on U . Namely, for each $1 \leq k \leq n$ and $k \leq j \leq n$, let $u_{k,j}$ solve

$$\begin{aligned} \Delta u_{k,j}(x) &= 0, \quad x \in U ; \\ u_{k,j}(x) &= c_{k,j}^i, \quad x \in \partial D_i \quad \text{and} \quad \int_{\partial D_i} \langle \nabla u_{k,j}(x), n(x) \rangle \nu_i(dx) = 0 \quad \text{for } 1 \leq i < k ; \\ u_{k,j}(x) &= 0, \quad x \in \partial D_i, \quad \text{for } i \geq k, i \neq j ; \\ u_{k,j}(x) &= 1, \quad x \in \partial D_j . \end{aligned}$$

The constants $c_{k,j}^i$ are not prescribed, i.e., solving the (k, j) -problem includes finding the boundary values of the function on $\partial D_i, i < k$. As in Section 2, the solution exists and is unique in $C^2(\overline{U})$.

Theorem 4.1. *Assume that $V_{k-1} < \lambda < V_k$ for some $1 \leq k \leq n$. Suppose that $T(\varepsilon)$ is such that $\lim_{\varepsilon \downarrow 0} (\varepsilon \ln T(\varepsilon)) = \lambda$. Let $\mathcal{E}_i, i = 1, \dots, n$, be arbitrary disjoint neighborhoods*

of O_i , $i = 1, \dots, n$. Then

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(X_{T(\varepsilon)}^{x,\varepsilon} \in \mathcal{E}_j) = u_{kj}(x), \quad x \in U, \quad j \geq k;$$

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(X_{T(\varepsilon)}^{x,\varepsilon} \in \mathcal{E}_j) = c_{k,j}^i, \quad x \in \overline{D}_i, \quad i < k, \quad j \geq k;$$

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(X_{T(\varepsilon)}^{x,\varepsilon} \in \mathcal{E}_j) = 1, \quad x \in \overline{D}_j, \quad j \geq k,$$

where $u_{k,j}$ is the solution of the (k,j) -problem and $c_{k,j}^i$ is the value of $u_{k,j}$ on ∂D_i . If $\lambda \geq V_n$, then

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(X_{T(\varepsilon)}^{x,\varepsilon} \in \mathcal{E}_n) = 1, \quad x \in \mathbb{T}^d.$$

Observe that $\sum_{j=k}^n u_{kj}(x) = 1$ for $x \in U$, while $\sum_{j=k}^n c_{k,j}^i = 1$ for $x \in \overline{D}_i$ if $i < k$. Thus, with probability close to one, $X_{T(\varepsilon)}^{x,\varepsilon}$ is located in a small neighborhood of one of the equilibrium points.

The proof of this theorem can be derived from the following facts by using the strong Markov property of the process.

1) For each $i < k$ and $x \in D_i$, the time it takes $X_t^{x,\varepsilon}$ to exit D_i is significantly smaller than $T(\varepsilon)$, i.e., $\tau^{x,\varepsilon}(\partial D_i)/T(\varepsilon) \rightarrow 0$ in probability as $\varepsilon \downarrow 0$.

2) For each $i \geq k$ and $x \in D_i$, the time it takes $X_t^{x,\varepsilon}$ to exit D_i is significantly larger than $T(\varepsilon)$, i.e., $\tau^{x,\varepsilon}(\partial D_i)/T(\varepsilon) \rightarrow \infty$ in probability as $\varepsilon \downarrow 0$.

3) The trace of the process $X_t^{x,\varepsilon}$ in \overline{U} converges to X_t^x as $\varepsilon \downarrow 0$.

We omit the details of the proof.

4.2 Trapping regions with multiple equilibriums

Up to now, we assumed that there was just one attractor (asymptotically stable equilibrium) inside each trapping region. Let us now consider an example where this is not the case. For simplicity, assume that there is one trapping region D containing two equilibriums O_1 and O_2 and one saddle point S . The structure of the vector v field on \overline{D} is assumed to be as shown in Figure 2. As before, the vector field is equal to zero in $U = \mathbb{T}^2 \setminus \overline{D}$.

Let $D_k \subset D$ and $\gamma_k \subset \partial D$, $k = 1, 2$, be the sets of points that are carried to an arbitrarily small neighborhood of O_k by the deterministic flow $\dot{x}(t) = v(x(t))$. Let $A, B \in \partial D$ be the points that separate γ_1 from γ_2 . Let γ be the curve that connects A with B and consists of two flow lines and the saddle point (see Figure 2). The asymptotic behavior of the process $X_t^{x,\varepsilon}$ (in exponential time scales) and of the trace $Y_t^{x,\varepsilon}$ is determined by the numbers V_{ij} defined in (2) and by the values $V_k = \inf_{x \in \partial D_k} V_k(x)$, where the quasi-potentials $V_k(x)$ are defined in (4).

Consider the case when $V_1 < V_{12}$, $V_2 < V_{21}$, and the infimum in the definition of V_k is achieved in a unique point $x_k \in \gamma_k$, $k = 1, 2$. Then, with probability that tends to one as $\varepsilon \downarrow 0$, the process exits D in an arbitrarily small neighborhood of x_k provided that it starts in a small neighborhood of O_k , $k = 1, 2$.

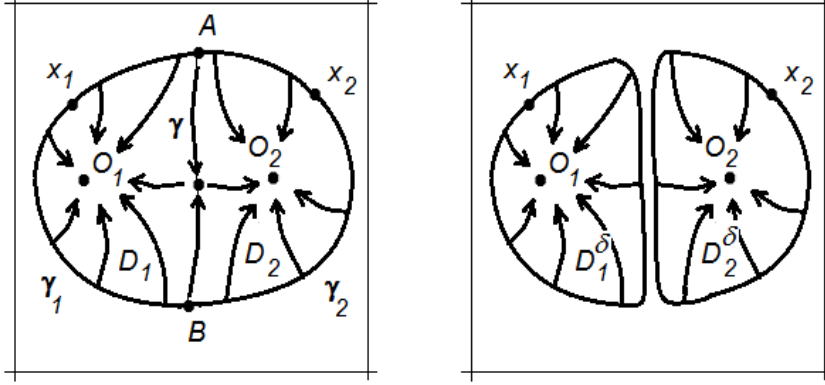


Figure 2: The flow lines of the original and the modified vector fields.

For $\delta > 0$, one can consider the following auxiliary system. Let V_δ be the δ -neighborhood of γ . Let $D_k^\delta \subset D_k$, $k = 1, 2$, be domains with smooth boundaries such that $D_k^\delta \setminus V_\delta = D_k \setminus V_\delta$, yet $\overline{D_k^\delta} \cap \gamma = \emptyset$. Moreover, we can modify the vector field v (i.e., replace it by a new vector field v^δ) in such a way that $v^\delta(x) = v(x)$ for $x \notin V_\delta$, while the field v^δ satisfies the assumptions with respect to the domains D_1^δ and D_2^δ that were imposed on v in Section 3, i.e., it is equal to zero outside $\overline{D_1^\delta} \cup \overline{D_2^\delta}$, is directed inside each of the domains on the boundary, and all the points of $\overline{D_k^\delta}$ are attracted to O_k .

The analysis of Section 3 applies to the process $X_t^{\delta,x,\varepsilon}$ defined via

$$dX_t^{\delta,x,\varepsilon} = \frac{1}{\varepsilon} v^\delta(X_t^{\delta,x,\varepsilon}) dt + dW_t, \quad X_0^{\delta,x,\varepsilon} = x.$$

Let $X_t^{\delta,x}$ denote the limit of the trace process in $\mathbb{T}^2 \setminus (D_1^\delta \cup D_2^\delta)$. Since $x_1, x_2 \notin V_\delta$ for all sufficiently small δ , it is not difficult to show that the probability that $X_t^{\delta,x}$ enters V_δ prior to time T tends to zero for each finite T . Moreover, there exists the limit in probability as $\delta \downarrow 0$ of $X_t^{\delta,x}$. This limit will be denoted by X_t^x . This process is the limit, as $\varepsilon \downarrow 0$ of the trace $Y_t^{x,\varepsilon}$ of the original process $X_t^{x,\varepsilon}$. A direct construction of the process X_t^x (in terms of the generator rather than via approximating processes) seems to be technically complicated and is not presented here.

Another case where the limit of the trace process can be easily described is when $V_1 < V_{12}$ if we assume that the infimum in the definition of V_1 is achieved in a unique point $x_1 \in \gamma_1$, while the infimum in the definition of V_2 is achieved in a unique point $x_2 \in \gamma \setminus \partial D$ (which implies that $V_2 = V_{21}$). Thus, with probability that tends to one as $\varepsilon \downarrow 0$, the process exits D in an arbitrarily small neighborhood of x_1 irrespective of whether it starts in D_1 or D_2 . The results of Section 3 then apply, with the limit of the exit measure μ being the point mass concentrated at x_1 .

A more general situation of several equilibria within D with various relations on

the quantities V_{ij} and V_k can be analyzed using the construction above based on removing the δ -neighborhoods of the boundaries of D_k and the results of Sections 6.5-6.6 of [4] on the hierarchies of cycles.

4.3 Other generalizations

If the process $X_t^{x,\varepsilon}$ is governed by a more general elliptic operator

$$L^\varepsilon = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{\varepsilon} v(x) \nabla, \quad (52)$$

then the results and the proofs are similar. The definition of the numbers V_{ij} and of the quasi-potential $V_k(x)$ should now be based on the action functional corresponding to the operator L^ε . The definition (8) of the measures ν_k needs to be modified to account for the variable diffusion coefficients of the process $X_t^{x,\varepsilon}$. However, if the infimum of $V_k(x)$, $x \in \partial D_k$, is achieved in a single point x_k , then ν_k is still the δ -measure concentrated at x_k .

The assumptions on the vector field v that we made in Section 3 do not specify that D necessarily contains a single equilibrium point. They may hold, for example, if D contains a single limit cycle instead. The case of several limit cycles is technically not different from the case of several equilibrium points that we discussed above.

The results also apply to processes on general smooth manifolds, not only on a torus.

4.4 More on the limiting process

The non-standard boundary problem introduced in Section 2 and the corresponding Markov process with jumps at the boundary arise in other situations, not just in the large deviation case. Consider, for example, a vector field v with closed flow lines that is equal to zero outside of \overline{D} such that ∂D serves as one of the flow lines (see Figure 3).

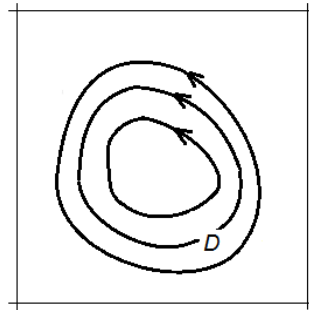


Figure 3: The flow lines inside D .

We expect that the trace in $\mathbb{T}^2 \setminus D$ of the process $X_t^{x,\varepsilon}$ with generator (52) converges, as $\varepsilon \downarrow 0$, to the process described in Section 2. The measure ν on ∂D will be defined by

the values of v in an arbitrarily small neighborhood of ∂D and by the diffusion coefficients a_{ij} on ∂D and can be calculated explicitly.

Acknowledgements: While working on this article, M. Freidlin was supported by NSF grant DMS-1411866 and L. Koralov was supported by NSF grant DMS-1309084.

References

- [1] Day M., *Mathematical Approach to the problem of noise induced exit*, Stochastic Analysis, Control, Optimization, and Applications. A volume in honor of W. H. Fleming. Editors: W. McEneaney, G. Yin, Q. Zhang. Birkhauser, Boston, 1999, pp 269-287.
- [2] Dynkin E. B., *Markov Processes*, Springer-Verlag, Berlin, Heidelberg, New York, 1965.
- [3] Freidlin M. I., *On Stochastic Perturbations of Systems with Rough Symmetry. Hierarchy of Markov Chains*, Journal of Statistical Physics, Vol. 157, No 6, pp 1031-1045, 2014.
- [4] Freidlin M. I., Wentzell A. D., *Random Perturbations of Dynamical Systems*, Springer 1998.
- [5] Mandl P., *Analytical Treatment of One-dimensional Markov Processes*, Springer-Verlag, 1968.